

teorema

Vol. XXX/2, 2011, pp. 79-94

ISSN: 0210-1602

[BIBLID 0210-1602 (2011) 30:2; pp. 79-94]

Of Boundaries and Metaphysical Starting Points: Why the Extended Mind Cannot Be So Lightly Dismissed

Joel Parthemore

RESUMEN

El debate en torno a la mente extendida versa, en gran medida, sobre dónde y cómo situar la frontera entre la mente cognitiva y el mundo no-cognitivo. Los conceptos de “interno” y “externo”, tomados del ámbito de los objetos físicos, se aplican como mucho metafóricamente a entidades como la mente. Las intuiciones defendidas enérgicamente revelan puntos de partida metafísicos que prejuzgan el debate. Si no se aceptan estos puntos de partida, las afirmaciones “sencillamente obvias” de los críticos de la mente extendida están mucho menos claras. La distinción entre mente y mundo es, en último término, una distinción conceptual, y como muchas si no todas las distinciones conceptuales, está sujeta a cambios a lo largo del tiempo.

PALABRAS CLAVE: *frontera, antirrealismo, realismo, intuiciones, enactivismo.*

ABSTRACT

The extended mind debate is, to large extent, a debate over where and how to locate the boundary between cognitive mind and non-cognitive world. Concepts of “internal” and “external”, taken from the domain of physical volumes, are metaphorical, at best, applied to entities like mind. Strongly held intuitions betray metaphysical starting points that prejudice the debate. If one does not accept these starting points, then the “just obvious” claims made by critics of extended mind are far less clear. The mind/world distinction is, ultimately, a conceptual one, and like most if not all conceptual distinctions, subject to shift over time.

KEYWORDS: *Boundary, Anti-Realism, Realism, Intuitions, Enactivism.*

I INTRODUCTION

Andy Clark and David Chalmers did not, of course, invent the notion of extended mind, versions of which can be found in enactive philosophy [Varela *et al.* (1991); Maturana and Varela (1992)], dynamic systems theory [Thelen and Smith (1994)], evolutionary psychology [Donald (1993)] and elsewhere. The idea has been around in one form or another for a long time

and has its roots, among other places, in the semantic externalism of Hilary Putnam [Putnam (1975)].¹ Nevertheless they have done more than anyone else to frame the contemporary debate and to show how various related lines of thought can usefully converge.

Clark and Chalmers open their now famous 1998 paper – a paper that has gone on to spawn something of a philosophical cottage industry – with a question: “Where does the mind stop and the rest of the world begin?” [Clark and Chalmers (1998), p. 7]. Although it is nowhere so well explicated as one might like, the concept of boundary is key. If *absolute fixed* boundaries of any kind are problematic – and I will argue that there is reason to think they are – then this one is particularly so. Like many boundaries, its precise location may seem straightforward, until we look too closely and start finding borderline cases. This paper will offer a distinctive argument that the boundary between mind and world (or, more or less equivalently, between self and other or between self and non-self) is pragmatic and conceptually defined in a way that, as I believe Clark and Chalmers correctly assess, shifts over time with respect to needs and perspective.

I will take “boundary” in this context – perhaps controversially – to mean a categorical dividing line such that specific instances of entities for whom the boundary is relevant *must* fall to one side of the boundary or the other.² Any entities that fall directly on the boundary indicate a problem at the least with where the boundary has been drawn, if not with the boundary itself. I will argue that issues of boundary deserve a great deal more attention than they have generally been given to date.

The argument will proceed as follows:

1. An examination of the current state of the extended mind debate shows that such key figures as Rob Rupert, Frederick Adams, and Kenneth Aizawa are relying on strongly held intuitions; and more importantly, they are making assumptions based on metaphysical starting points that they are less than straightforward about. To some extent Andy Clark and Dave Chalmers are as well, though Clark in particular is careful about hedging his bets.
2. If one does not accept these metaphysical starting positions, then the “just obvious” nature of many of the claims made by critics of extended mind is much less obvious. In particular, their metaphysics predispose them to take concepts of inner and outer (internal/external) from the physical domain and apply them to cognition, where it might appear that such application is loosely metaphorical at best.

3. There are reasons independent of metaphysics for questioning the application of concepts like “internal” and “external” to the cognitive domain. However, if one allows anti-realism as at least a plausible position for sake of argument, then those concerns will be heightened.
4. Concepts provide a particularly fruitful area to focus on here, because they give us reason to believe that our concepts and conceptual prejudices affect all that we experience. There are sound logical reasons to believe that concepts and their referents cannot fully pull apart, relative to the perspective of any conceptual agent (who cannot step outside his own conceptual frame of reference).
5. If concepts bleed into the world in this way, and if concepts are intrinsically cognitive entities, then cognition meaningfully extends into the world as well. Bottom line: the distinction between mind and world is a conceptual distinction that cannot reliably be separated from any ontologically prior one (if there even is any).

In *Supersizing the Mind* [2008], Clark talks about the ways in which the boundary we draw between the self and the non-self (self and other, self and environment) changes over time (in particular for any given individual, but perhaps collectively as well). This is to be expected if 3-4 are true and is, I believe, the best way of advancing a version of the extended mind hypothesis. Yes, there are things that are clearly and untendentiously “self” and things that are clearly and untendentiously “non-self” (at least at any given point in time, but often in a stable way over time as well). At the same time there are other, borderline cases that appear to fall on either one side of the boundary or the other depending on one’s perspective at the time. The “inner”/“outer” “internal”/“external” distinction with respect to cognition is deeply problematic, precisely because of this difficulty with boundaries.

To wit, and *pace* Rupert, and Adams and Aizawa, some version of the extended mind hypothesis belongs – indeed, *must be*, if I am right – on the table for discussion.

However, before proceeding to this positive account for advancing the extended mind hypothesis, I wish to address how I think the extended mind hypothesis should *not* be advanced. The rest of the paper is structured as follows. Section II addresses difficulties with the familiar Otto thought experiment and raises a common worry about *cognitive bloat*. Section III presents my best understanding of the current state of the extended mind debate and raises questions about the driving intuitions and the metaphorical starting points of the various players. Metaphysical concerns aside, strict boundaries just are, from a conceptual point of view, problematic entities, and certain boundaries (as between mind and world) even more so. Section IV summa-

rizes the overall argument once more and offers some proposals for moving the extended mind discussion forward.

II. WHAT'S WRONG WITH OTTO?

Clark and Chalmers introduce their readers to Otto and Inga, two New York residents who, for sake of the thought experiment, share a common belief (the location of the Museum of Modern Art) and a common desire (the intention to go there for a certain exhibition). For both of them the belief is meant to be *non-occurrent*: so for example, Inga is not consciously thinking about the location of the museum before she hears about the exhibition, but we would still conventionally ascribe to her the belief that the museum is on 53rd Street, based on her prior acquisition of that knowledge. Otto, also, is not consciously thinking about the location of the museum before he hears about the exhibition, but that is because, as an Alzheimer's patient with significant memory impairment, he cannot. Instead he has written the museum's location in the notebook that he carries everywhere with him, and where he records all information that might at some point prove important.

The only relevant difference between Otto and Inga is meant to be the location of the information, prior to conscious recall: conventional long-term memory for Inga, notebook for Otto. If Inga has a non-occurrent belief, then so must Otto – or so the reasoning goes.

I find Otto and Inga a distraction for two reasons. The first is its real-life implausibility; but, of course, the concerns raised by the thought experiment may be valid regardless. More substantially, one might reasonably worry that Clark and Chalmers fail to pick out what is distinctive about hypothetical Otto: namely, the very *intimate* way he is bound to his notebook, the consulting of which seems to be the one thing he reliably can remember. He *cannot* consult his non-notebook-based memory of the museum's location on some semi-reliable basis, because then the notebook would be optional, and that would violate Clark's condition that "the resource be reliably available and typically invoked" [Clark (2008), p. 79].³

On face value, and *pace* Chalmers' discussion about his iPhone in the forward to *Supersizing the Mind*, most people simply are not dependent on their notebooks, laptops, mobile phones, and other such props, in anything like the same way. If Otto lost his notebook, he would, in a very real way, lose a part of himself: that seems plausible. In contrast, when the hard drive on my laptop crashes, I am generally and genuinely put out, but I am hardly transformed into a different person.

I am not rejecting the extended mind hypothesis – far from it! I am rather questioning the choice of supporting examples and, to some extent, the method of argument. Yes, Otto and his notebook would, if they existed, con-

stitute a *prima facie* plausible candidate for an instance of extended mind; however, the example does not generalize in the way Clark and Chalmers want or, indeed, need it to. The extended mind hypothesis is a lot less interesting if it is limited to a few rare if not, indeed, merely hypothetical, cases.

Worse, if it did generalize that way, issues of cognitive bloat raise their head. Clark would do well to remember where the term “supersizing” comes from: the American fast food industry; and what connotations the term had and has there. The moral of the story may be that a little extended mind goes a long way.

III. IN SEARCH OF A BETTER ARGUMENT FOR EXTENDED MIND

Supersizing the Mind is in many ways an extended re-statement of the main ideas from the 1998 paper and response to the critics. It summarizes the extended mind hypothesis like this:

Proponents of the extended mind story hold that even quite familiar human mental states (e.g., states of believing that so and so) can be realized, in part, by structures and processes located outside the human head. Such claims go far beyond the important but far less challenging assertion that human cognizing leans heavily on various forms of external scaffolding and support. Instead, they paint mind itself (or better, the physical machinery that realizes some of our cognitive processes and mental states) as, under humanly attainable conditions, extending beyond the bounds of skin and skull [Clark (2008), p. 76].

The main points I take to be these:

- Any version of mind/brain identity [e.g., Churchland (1989)] is rejected. Mind is neither the same as nor reducible to brain; and in particular, they need not share the same boundary with respect to the world.
- It is important, but insufficient, to stress the rich interactions between any cognitive agent and its environment. At least some of those interactions are sufficiently rich so as to blur the lines between the two.

It is telling that Clark refers to “scaffolding”. It is in the nature of scaffolding that it can be removed once the structure it is supporting is complete. The structures Clark believes to contribute to instances of extended cognition are *not* scaffolding precisely because they cannot be removed (at least, not without the cognitive agent becoming something substantially different from what

it was). They become, in some non-trivial sense, if Clark is right, part of the cognitive agent.

III.1 *Intuitions and Counter-Intuitions*

Although considerable empirical evidence is cited, much of the argument for the extended mind hypothesis comes from an intuition pump commonly referred to as The Parity Principle (though it is not called that in the original paper): “if, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process” [Clark and Chalmers (1998), p. 8].

Criticism of the Clark-Chalmers position has arisen most prominently from, on the one hand, Fred Adams and Ken Aizawa, who published a seminal paper in 2001 [Adams and Aizawa (2001)]; and on the other, Robert Rupert, who published a similarly influential article in 2004 [Rupert (2004)]. Borrowing a page from Clark, both have recently published books along the same lines as their earlier papers: [Adams and Aizawa (2008)] and [Rupert (2009a)]. In what follows I will make the most reference to Rupert’s book, which is the more recent and makes many of the same arguments.

Adams, Aizawa and Rupert all claim sympathy for the Parity Principle. “To us, [the Parity Principle] means that the skull does not constitute a theoretically significant boundary for cognitive science. More specifically, it means that being inside the brain cannot be the mark of the cognitive. This seems to us true and obvious” [Adams and Aizawa (2001), p. 46]. “I sympathize with the motivation behind the Parity Principle. After all, why should it matter where a process takes place? If that process instantiates cognitive or mental properties when it is over *here*, why should things change simply because it is now over *there*?” [Rupert (2009a), p. 30].

All claim, however, a powerful *counter*-intuition. “...To many in the philosophical community, the extended view seems incredible on its face” [Rupert (2009b), p. 314]. Aizawa is more colorful when he describes, in the introduction to *The Bounds of Cognition*, his introduction to the extended mind. “Some time in the early summer of 1998 or so, Fred came across a paper by Andy Clark and David Chalmers, advancing what seemed to us to be the outrageous hypothesis that, at least at times, cognitive processes extend into the tools people use” [Adams and Aizawa (2008), p. vii].

Regardless – whatever value intuitions may have initially, they are – on Rupert’s, or Adam’s and Aizawa’s, account – things ultimately to be set aside, like idiosyncrasies of perspective. “The average person’s intuition-based applications of ‘cognition’, even the well-informed theorist’s reactions, should not be trusted to reflect the *actual* structure of cognition – unless, of course, the subject has in hand the correct theory of cognition” [Rupert (2009a), p. 32]. “After all, it’s not up to our intuitive judgments to decide

what cognition is; the property of being cognitive is a scientific construct, validated only by the causal-explanatory work it does” [Rupert (2009b), p. 323].

All three insist that the limitation of the mind to the boundaries of the body, if not the brain, is contingent and not *a priori*: i.e., no “fetish for the bodily boundary” [Rupert (2009a), p. 45]. To wit, and *contra* Clark and Chalmers’ reading, they claim that the extended mind hypothesis, in any meaningful form, flies in the face of nearly all empirical evidence, and that all of the explanatory work that the extended hypothesis might claim to do can be done as well or better simply by emphasizing the agent’s *embodiment* (the agent takes a particular physical form, by which it interacts with its environment) and *embeddedness* (the agent is situated in a particular environment, with which it richly interacts).

Perhaps the biggest concern, which Adams, Aizawa, and Rupert again all share, is that – ironically, given the title of Clark’s book – the extended mind hypothesis will lead to inevitable cognitive bloat, whose “...threat is of pancognitivism, where everything is cognitive. This is surely false” [Adams and Aizawa (2001), p. 57]. “Does the explanatory principle [on which the extended mind hypothesis might be grounded] entail that my house’s state of being structurally unsound is partly located where the builders’ corpses are? Or that the past mental states of these now deceased people are part of the physical substrate my house’s property of being structurally unsound?” [Rupert (2009a), p. 20] Adams and Aizawa offer similar *reductios ad absurdum*.

Arguments back and forth have led to a succession of papers and books. It does not seem to me, however, that any of the principal parties to this debate have significantly moved on or changed their position. Instead they have often resorted to legalistic language reminiscent of lawyers arguing an obscure point of corporate law. Once one clears the legalese away, the differences – somewhat of emphasis, somewhat of substance – are fairly straightforward; and, unlike obscure matters of corporate law, they really do (and should!) matter to the rest of us. After all, the boundary we draw between self and non-self is, plausibly, foundational to all our other conceptual distinctions.

III.2 *Why Metaphysics Matters*

It will be useful to begin a discussion of metaphysical premises by saying what the metaphysical disagreement is not, which is a debate about materialism or naturalism, or about Cartesian dualism versus physical monism. All of the parties to this debate, with the notable exception of Chalmers⁴, are keen to stress their materialist credentials: e.g., “In questioning BRAINBOUND⁵, I shall not in any way be questioning the basic materialist vision of mind as emerging fully and without residue from physical goings-on” [Clark (2008), p. xxviii].

So what is the disagreement about? Rupert puts a lot of weight on the word “literal”: the word is literally peppered throughout his book. In a typical passage he describes the extended view as “the view that human cognition – to

some substantial degree – literally includes elements beyond the boundary of the human organism” [Rupert (2009a), p. 3]. The implication, I assume, is that Clark and Chalmers are not “merely” speaking metaphorically: they “really” mean that. Such a crisp literal/metaphorical distinction – whether with one’s language or one’s concepts – assumes some form of realist metaphysics, as is clear from many places in Rupert’s writings (though nowhere have I seen it stated so baldly). The literal meaning is the fact of the matter that realism aims to deliver. But there is nothing about physicalism, so far as I am aware, that entails realism.

On this point Clark is holding his cards in his hands, while Chalmers is, I believe, best understood as both a physicalist (he rejects substance dualism) and an anti-realist (he takes experience as foundational: i.e., something that must be assumed from the beginning and cannot be subtracted out) [Chalmers (1996)]. Contrast Rupert with Peter Gärdenfors, whose avowedly anti-realist leanings inform his theories about concepts: “the upshot is that [in conceptual spaces theory] there is no sharp distinction between literal meaning and metaphor” [Gärdenfors (2004), p. 187].

Realism I take as the metaphysical assumption either (*per* direct realism) that the apparent transparency of the world should, in most instances at least, be taken at face value; or (*per* indirect realism) that if the apparent transparency cannot be taken at face value, it can, at least, be logically reconstructed. In either case, science talks about the world in a perspective-free (or essentially perspective-free) way. Anti-realism is the position that, while the fully mind-independent world is conceded logically to exist, one cannot, as a matter of principle, say anything about it, beyond its bare existence and its ongoing role in constraining experience. Put another way: what Husserl called the *lifeworld* is always, in some way and to some ineliminable extent, touched by mind.

Let me be clear: anti-realism is not the perspective that world *is* mind; that would be idealism. Neither does anti-realism allow one to believe whatever one likes about the world: if the world constantly outruns our conceptual understanding of it, at the same time it constantly and forcibly constrains that understanding, sometimes on pain of injury or death.

Anti-realism, pragmatism, and pluralism go hand in hand, where pragmatism/pluralism is taken as the position that there need be, in most instances at least (and perhaps all of import), no single fact of the matter. Pragmatism can tolerate apparent contradictions, so long as they are qualified by perspective: e.g., p from one perspective, $\sim p$ from another. So long as one does not try to hold both perspectives at once – i.e., make them part of a single unified perspective – there is no contradiction in defending both.

In this light, Rupert’s statement that “even if one is inclined toward pluralism, an extended and an embedded model cannot both be true of a single cognitive process – else there is a single cognitive system that both extends

beyond the boundary of the organism and does not” [Rupert (2009a), p. 9] is, on the face of it, simply wrong: whether it is seen to extend or not may simply be, on a pluralist view, a matter of the perspective one is taking, the questions one is asking, and the intended purpose to which one will put the answers. Pluralism comes with no guarantees that our conflicting perspectives can necessarily be reconciled into a single unified perspective.

I am arguing here neither for the correctness of anti-realism nor the falsehood of realism. It is enough to allow a modest anti-realism as a plausible position for sake of argument. If some form of anti-realism *should happen* to be true, then intuitions, like perspectives, cannot simply be set aside; they will play an unavoidable and substantive role in the theory.

III.3 *Metaphysical Premises and Boundaries*

Metaphysical premises become clearest when one looks at another word used ubiquitously by all parties to this debate, and the one with which I began this paper: namely, “boundary”. After all, the extended mind debate at heart is about where one should draw the boundary between mind and world, and whether that boundary is fixed at the physical boundary of “skin and skull”. For all of the importance (rightly) placed on this boundary, one might expect there to be more attention paid not just to locating it correctly but determining its nature. Is the boundary “really” real, or is it something we construct (and can move)?

Overly rigid boundaries of any kind can prove problematic, if one probes them too closely. Consider cell boundaries, as clear a boundary as one is likely to find. Any effective cell membrane must be permeable: a continuity to match the discontinuity. The problem is: at what precise point does a molecule pass from “outside” to “inside”? The closer one examines the cell boundary, the harder the answer becomes. The answer is only clear if one observes from a sufficiently detached perspective.

Boundaries at the level of multicellular organisms only become more difficult. Is my bodily boundary at my epidermis (layer of dead skin cells) or my dermis (live cells)? It depends upon the context in which you’re asking. Likewise my body is, in terms of topology theory, torus shaped, like a donut with a hole through the middle (as opposed to a filled donut). Normally I think of my digestive tract as ‘inside’, but, as with the empty space in the middle of the ring donut, it is, from some perspective, ‘outside’. Even though the ring donut defines the empty space in the middle, the empty space is not actually part of the donut.

What of the bacteria living in my gut, who depend on me for their existence, and whom I likewise depend on for mine? Are they inside or out? Are they part of me or not? I am reliably informed that several kilograms of my body weight consist of single-celled organisms. When I weigh myself in the morning, I do not mentally subtract them.

The realist, of whatever persuasion, need not, of course, be bothered about any of this. *Prima facie*, it is enough for her to say e.g. that the boundary between mind and world is *roughly* at the physical boundary of skin and skull – or is it?

The difficulty (or, for the extended mind enthusiast, the opportunity) lies with how rough is “roughly”, and in particular with the way Rupert (along with Adams and Aizawa) moves seamlessly from the boundary of the organism *as a biological agent* to the boundary of the organism *as a cognitive agent*. “Internal” and “external” are attributes of physical volumes, and it is not immediately clear that a cognitive agent is that sort of thing. Yet Adams and Aizawa write, “To ask about the bounds of cognition is to ask what portions of spacetime contain cognitive processing.... It is to ask about the physical substrate of cognition” [(2008), p. 16].

This is, perhaps, not a problem, provided one sees mind reducing to brain *per* a reductionist or eliminativist account, or emerging from brain in a way that is either immediately transparent or reconstructably so. (The latter route is, I think, the one that Rupert wants to take: he seems ready to allow that mind *could* just be a functional description with no immediate physical translation.) Either mental boundaries “just are” physical boundaries, or they map straightforwardly to them.

Reconstructable in principle, however, need not mean reconstructable in practice; and herein lies the fruitful middle ground between anti-realist and realist perspectives: without that reconstructability, and on closer examination, mental boundaries look woefully unclear. At the same time, a clear and at least relatively fixed mental boundary is essential to Rupert’s arguments. Contrast this with Clark’s flexible notion of the same boundary, when applied to what he terms “profoundly embodied agents”: “such agents are able constantly to negotiate and renegotiate the agent-world boundary itself. Although our own capacity for such renegotiation is, I believe, vastly underappreciated, it really should come as no great surprise, given the facts of biological bodily growth and change” [(2008), p. 34].

In keeping with much if not most of the literature in this field, Rupert talks about representations without defining what they are: the assumption is that the definition is already understood and agreed upon. Unfortunately, proceeding to label some representations as “internal representations”, as Rupert does, does not help unless the application of “internal” in the mental domain is *also* already understood and agreed upon. I have offered reasons to think it is not. Rupert’s offer of a “systems-based criterion”⁶ is, on its own, no help in recapturing a clear sense of boundary unless that criterion is assuming the very physical translation that is meant to be derived – no help, certainly, if one gives any weight to concerns like this one from Clark:

Nontrivial causal spread... occurs whenever something we might have expected to be achieved by a certain well-demarcated system turns out to involve the exploitation of more far-flung factors and forces [(2008), p. 7].

One school of thought actively trying to occupy the middle ground is enactivism, as defined by Francesco Varela: “I have proposed using the term *enactive* to... evoke the idea that what is known is brought forth, in contraposition to the more classical views of either cognitivism or connectionism” [Maturana and Varela (1992), p. 255]. In common with extended mind, enactivism – as used by Varela, Evan Thompson, John Stewart and others [Varela *et al.* (1991); Thompson (2007); Stewart (1995)] – views cognition as spanning brain, body and environment (to paraphrase a phrase used by Clark and by many, many others). Like extended mind, enactivism includes but goes beyond notions of embeddedness and embodiment by:

- Understanding cognition, at least in the first instance, as a *skillful activity*;
- Emphasizing the ineliminable role of an observer, which is to say, a first-person perspective⁷;
- Stressing a continuity between agent and environment as underlying any conceptual distinctions between the two⁸;
- Consequently presenting agent and environment as co-emergent.

Although not explicitly enactive, Gärdenfors’ conceptual spaces theory fits comfortably into such a perspective⁹. Clark, pointedly, has not endorsed enactivism. But perhaps he should.

III.4 The Argument from Concepts

One of the stock phrases in [Adams and Aizawa (2001)] and [Adams and Aizawa (2008)] is the “mark of the cognitive”, the hallmark of which is “non-derived content”. Regardless of what else does or does not bear the “mark of the cognitive”, Adams and Aizawa must surely agree that concepts, as the building blocks of structured thought, do. If concepts are not mental entities, what are? And if they are not mental entities, what are they?

It is a trivial observation, almost a definitional truism, that our structured understandings of the world are conceptually mediated. The interesting question is whether or not concepts are transparent to the pre- or non-conceptual world (either immediately or reconstructably so). What is the relationship between clearly conceptual mind and seemingly non-conceptual world? To wit, is there a pre- or non-conceptual experiential Given that then rationally justifies our conceptually structured beliefs and experience? John McDowell, borrowing a page from Wilfrid Sellars [Sellars (1956)], calls this

“the Myth of the Given”: the idea that “experience, conceived in such a way that it could not be a tribunal, is nevertheless supposed to stand in judgment over our empirical thinking” [(1996), p. xvii].

McDowell is likewise famous for saying that concepts go “all the way out”. Though he is generally careful to qualify where they go “all the way out to” – e.g., “all the way out to the world’s impacts on our receptive capacities” [McDowell (2007), p. 338] – nonetheless, the practical consequence is generally taken to be that the world, itself, is somehow (fully) conceptual. So Jeremy Koons reads McDowell’s attack on the Myth of the Given as meaning that “world is assimilated to mind: reality is in the space of reasons. Mind is able to represent world because world, like mind, turns out to be conceptual: reality is itself conceptual” [Koons (2004), p. 130].

One need not embrace McDowell’s strong conceptualism – the position that the content of experience is *fully* conceptual, fully within what he calls the “space of reasons” – in order to accept his attack on the Myth of the Given. (I, for one, do not.) So long as concepts do not fully or reliably pull apart from their referents in the world, so long as some conceptual residual remains (or even may remain), then the conceptually untainted Given will be a myth. More importantly for present purposes, one need not embrace any degree of anti-realism to allow the possibility, in practice, of such residue. If it exists, even potentially, in the *lifeworld*, and it bears the “mark of the cognitive”, then, in a “literal” sense, cognition exists in (extends into) the world as well.

Why might one think that such residue exists? Consider: there are concepts, and there are things that are not concepts, with a *prima facie* clear boundary between, as clear as the boundary appears between mental self and non-mental world. Yet to reflect upon either (as one must, in order to discuss them) is to bring them into the space of concepts: so now one has concepts of concepts (higher-order) on the one hand, concepts of non-concepts (first-order) on the other. Meanwhile *that* reflection invites another layer of regress, and so on. To reject this insight and arbitrarily break the regress is to invite paradox.

Consider your pet dog Fella. Upon any specific Fella encounter, you bring a great deal of conceptual expectations to bear, whether you are reflectively aware of doing so or not (and, in the usual circumstances, you probably will not be). What you experience is an object, with all the expectations of e.g. object permanence that you have had from a very early age; but not just any object: a dog, with all your expectations about dogs (that *ceteris paribus* they *do* bark, that they *never* purr or meow); but not just any dog: your dog Fella, whom you’ve raised since he was a puppy, who crawls into your bed every night, who you took to the vet’s last Tuesday for de-worming. Whatever it is, the risk if not the certainty remains that the referent of “my dog Fella” is not, or is never just, the thing-in-itself, stripped of all conceptual shading. Even the concept “my dog Fella yesterday” is a generalization over,

or abstraction away from, many specific Fella moments over the course of the day. To recognize Fella *as Fella* – or, more minimally, as a dog, or more minimally yet, as a stable and re-identifiable object – is already to have passed beyond the possibility of any strictly in-the-moment, strictly non-conceptual experience.

All this is to be expected if one considers that the distinction between conceptual and non-conceptual, like the distinction between mind and world, is a conceptual and not an ontologically prior one. Unless one assumes some type of *natural kinds* argument, there is no *prima facie* reason to think that any conceptual distinctions match one-to-one with category boundaries in the world, or that the latter concept is necessarily coherent. Rupert acknowledges the conceptual distinction and its proneness to shift (see e.g. the discussion in Rupert [(2009a), p. 166], but says the boundary he is interested in is the *real* one. I claim that, at the least, the conceptual boundary cannot reliably be separated from any ontologically prior one. Meanwhile, as concepts change, so do the boundaries they draw: not *too* much, or the conceptual structure breaks down; but *just enough* for them to continue to be relevant to present circumstances.

IV. CONCLUSIONS

I can now return to the question with which I (and Clark and Chalmers) began: “Where does the mind stop, and the rest of the world begin?” After all, one has to draw the line somewhere: otherwise, if everything is (equally) cognitive, the term loses any usefulness.

As should be clear, I share many of Rupert’s, and Adam’s and Aizawa’s, concerns about cognitive bloat. When one “supersizes” the mind, one should be careful not to go too far. Even if the keyboard might, in some circumstances, become part of my cognitive system, it should not do so too easily.

Besides the Parity Principle, the most well-known intuition pump for the extended mind hypothesis is the Otto-Inga thought experiment. If it works, I have argued, it is only because Otto is a very rare individual, in a way that Clark and Chalmers don’t properly address. As Rupert writes, “even if there are some cases like Otto’s, they seem too rare to drive a paradigm shift in cognitive science” [(2009a), p. 90]. I have agreed with Rupert that the alternative possibility – that Otto represents something in fact quite common – is worse.

The moral of the story is that a little extended mind goes a long way; and all that this requires is a sufficiently flexible boundary between mind and world: one that (as a conceptually defined boundary) shifts not too much but *just enough*. An over-concern with the “literal” truth of the matter reflects metaphysical agendas. Metaphysical premises may lead one to see that bound-

ary as more rigid than experience suggests it is, and to downplay experienced reality in favor of the “literal” facts of the matter.

If concepts bleed into the *lifeworld*, and if concepts are intrinsically cognitive entities, then cognition plausibly extends some indeterminable (and variable) distance into the world as well. One need not maintain – as McDowell is often read – that the world we encounter is *fully* conceptual, which would lead to cognitive bloat; only that there is no part of that world that is fully or reliably free of the conceptual touch. Such cognitive tentacles into the world are all that the extended mind hypothesis, on my reading, requires.

It seems to me that discussion on extended mind could best be moved forward not by debating what the *one* correct perspective is – extended, or merely embodied/embedded – but by a renewed search for practical explanatory benefits in stressing the continuity between agent and environment. Here, I think, the greatest benefit may be in shifting away from the default assumption that *physically* instantiated phenomena are always conveniently *localized* entities with clearly defined boundaries. If any specific variant on physicalism should be called into question, it is this one.

Centre for Cognitive Semiotics
University of Lund
Box 201, 22100 Lund, Sweden
E-mail: joel.parthemore@semiotik.lu.se

NOTES

¹ Clark and Chalmers are, however, careful to distinguish their *active externalism* from Putnam’s own, passive externalism.

² Of course, many useful boundaries in other domains are not so sharply drawn, as one of the reviewers pointed out. However, I take the debate over the extended mind hypothesis to revolve, in large part, on whether the mind/world boundary *can* be so sharply drawn. If that boundary is sufficiently flexible, then Clark and Chalmers have made their point.

³ In case of doubt, Clark adds, tellingly, that “Otto always carries the notebook and won’t answer that he ‘doesn’t know’ until after he has consulted it.”

⁴ Chalmers is an interesting case. Although subject to frequent claims of being a Cartesian dualist, Chalmers is avowedly *not* one (see e.g. [Chalmers (1996)], preferring what he calls there (and elsewhere) a more “innocent” dualism that is meant to be fully compatible with an orthodox scientific world view and physical theory.

⁵ BRAINBOUND is Clark’s name for the position he contrasts with the extended view.

⁶ “...I argue that the relatively durable cognitive system – the integrated collection of capacities and mechanisms that causally contributes to the production of cognitive phenomena – provides the most plausible line of demarcation between what

is cognitive and what is not.... Given that this integrated system typically appears within the boundary of the human organism, cognition does not extend beyond that boundary, at least not in the substantive way supposed to lead to paradigm shift in cognitive science" [Rupert, 2009, p. 7].

⁷ "Everything that is said, is said by an observer to another observer that could be himself" [Maturana, 1978, p. 30].

⁸ Such an emphasis on continuity might on first blush seem at odds with the central importance and seeming inflexibility of boundaries – notably the cell boundary – to Maturana and Varela's notion of autopoiesis (see for example [Maturana and Varela, 1992, 1980]). This, I think, would be a severe misreading of Maturana and Varela. The membrane is essential, yes, but only relative to one's perspective as an observer; it is absolutely critical to their notion of autopoiesis that, from some other perspective, what matters is the *continuity* between organism and environment, each actively defining the other. That is to say, the boundary is only meaningful with respect to that continuity.

⁹ As Gärdenfors himself acknowledges (personal communication).

REFERENCES

- ADAMS, F. and AIZAWA, K. (2001), 'The Bounds of Cognition', *Philosophical Psychology*, vol. 14, pp. 43-64.
- (2008), *The Bounds of Cognition*, Chichester, John Wiley and Sons.
- BALLARGEON, R. (1987). 'Object Permanence in 3 1/2- and 4 1/2-Month-Old Infants', *Developmental Psychology*, vol. 23, pp. 655-664.
- CHALMERS, D. J. (1996). 'Facing Up to the Hard Problem of Consciousness', in Hameroff, S. R., Kaszniak, A. W., and Scott, A. (eds.), *Toward a Science of Consciousness: the First Tucson Discussions and Debates*, Cambridge, Mass., MIT Press, pp. 5-28.
- CHURCHLAND, P. S. (1989), *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge, Mass., MIT Press.
- CLARK, A. (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, Oxford, Oxford University Press. Also available as ebook through Oxford University Press (<http://www.oxfordscholarship.com>).
- CLARK, A. and CHALMERS, D. (1998), 'The Extended Mind', *Analysis*, vol. 58, pp. 7-19.
- DONALD, M. (1993), *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*, Cambridge, Mass., Harvard University Press.
- GÄRDENFORS, P. (2004), *Conceptual Spaces: The Geometry of Thought*, Cambridge, Mass., Bradford Books. First published in 2000.
- KOONS, J. R. (2004), 'Disenchanted the World: McDowell, Sellars, and Rational Constraint by Perception', *Journal of Philosophical Research*, vol. 29, pp. 125-152.
- MATURANA, H. (1978), 'Cognition', in Hejl, P. M., Köck, W. K., and Roth, G. (eds.), *Wahrnehmung und Kommunikation*, Frankfurt, Peter Lang, pp. 29-49. Available online at <http://www.enolagaia.com/M78bCog.html>, with the original page numbering retained.
- MATURANA, H. and VARELA, F. J. (1980), *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Kluwer Academic Publishers.
- (1992), *The Tree of Knowledge: The Biological Roots of Human Understanding*, London, Shambhala.

- MCDOWELL, J. (1996), *Mind and World*, Cambridge, Mass., Harvard University Press.
- (2007), ‘What myth?’ *Inquiry*, vol. 50, pp. 338-351.
- PIAGET, J. (1954), *The Construction of Reality in the Child*, New York, Basic Books.
- PUTNAM, H. (1975), ‘The Meaning of ‘Meaning’’, in Gunderson, K. (ed.), *Language, Mind, and Knowledge*, Minneapolis, University of Minnesota Press.
- RUPERT, R. (2004), ‘Challenges to the Hypothesis of Extended Cognition’, *The Journal of Philosophy*, vol. 101, pp. 389-428.
- (2009a), *Cognitive Systems and the Extended Mind*, Oxford, Oxford University Press.
- (2009b), ‘Critical Study of Andy Clark’s *Supersizing the Mind*’, *Journal of Mind and Behavior*, vol. 30, pp. 313-330.
- SELLARS, W. (1956), ‘Empiricism and the Philosophy of Mind’, in Feigl, H. and Scriven, M. (eds.), *Minnesota Studies in the Philosophy of Science, Volume I: The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, Minneapolis, University of Minnesota Press, pp. 253-329. Available online from <http://www.ditext.com/sellars/epm.html>.
- STEWART, J. (1995), ‘Cognition = Life: Implications for Higher-Level Cognition’, *Behavioural Processes*, vol. 35, pp. 311-326.
- THELEN, E. and SMITH, L. (1994), *A Dynamic Systems Approach to the Development of Cognition and Action*, Cambridge, Mass., MIT Press.
- THOMPSON, E. (2007), *Mind in Life: Biology, Phenomenology and the Sciences of Mind*, Cambridge, Mass., Harvard University Press.
- VARELA, F. J. and MATURANA, H. (1992), *The Tree of Knowledge: The Biological Roots of Human Understanding*, London, Shambhala.
- VARELA, F. J., THOMPSON, E., and ROSCH, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, Mass., MIT Press.